Time series clustering to examine presence of decrement in Parkinson's finger-tapping bradykinesia

Conference Paper · July 2020

DOI: 10.1109/EMBC44109.2020.9175638

7 authors, including:

Zhibin Zhao
Xi'an Jiaotong University
44 PUBLICATIONS **281** CITATIONS
SEE PROFILE

Hui Fang
Loughborough University
49 PUBLICATIONS **305** CITATIONS
SEE PROFILE

Jane E Alty
University of Tasmania
81 PUBLICATIONS **506** CITATIONS
SEE PROFILE

David Wong
The University of Manchester
34 PUBLICATIONS **242** CITATIONS
SEE PROFILE

Time series clustering to examine presence of decrement in Parkinson's finger-tapping bradykinesia

Zhibin Zhao[1], Hui Fang[2], Stefan Williams[3], Samuel D. Relton[3], Jane Alty[4], Alex J. Casson[5], David C. Wong[6]

***Abstract*— Parkinson's disease is diagnosed based on expert clinical observation of movements. One important clinical feature is decrement, whereby the range of finger motion decreases over the course of the observation. This decrement has been assumed to be linear but has not been examined closely.**

We previously developed a method to extract a time series representation of a finger-tapping clinical test from 137 smartphone video recordings. Here, we show how the signal can be processed to visualize archetypal progression of decrement. We use k-means with features derived from dynamic time warping to compare similarity of time series. To generate the archetypal time series corresponding to each cluster, we apply both a simple arithmetic mean, and dynamic time warping barycenter averaging to the time series belonging to each cluster.

Visual inspection of the cluster-average time series showed two main trends. These corresponded well with participants with no bradykinesia and participants with severe bradykinesia. The visualizations support the concept that decrement tends to present as a linear decrease in range of motion over time.

***Clinical relevance*— Our work visually presents the archetypal types of bradykinesia amplitude decrement, as seen in the Parkinson's finger-tapping test. We found two main patterns, one corresponding to no bradykinesia, and the other showing linear decrement over time.**

I. INTRODUCTION

Bradykinesia is the core clinical feature of Parkinson's disease. It is defined as a pathological slowness of movement and decrement in amplitude or speed of movement (or progressive hesitations/halts) as movements are continued [1]. It is assessed by a specialist clinician observing the patient repetitively tapping index finger and thumb together as 'wide and fast as possible' over ten taps or ten seconds. The concept of decrement is key to the clinical distinction between Parkinson's bradykinesia and slow but stable movement seen in normal ageing or other conditions such as arthritis.

Decrement is currently described within two clinically validated scales. In the Unified Parkinson's Disease Rating Scale (UPDRS), it is categorized by how soon in the test the reduction in amplitude begins, but no minimum reduction is defined [2]. In the Modified Bradykinesia Rating Scale (MBRS), amplitude (which includes decrement) is defined on a scale of 0 (normal) to 4 (severe). For example, a score of 1 occurs when 'Mild reduction in amplitude in later performance, most movements close to normal'[3].

Previous attempts to precisely quantify decrement in the clinical literature have assumed linear decrease in amplitude over time. For instance, both Bank et al. and Martinez-Manzanera et al. quantify decrement by computing a linear regression of the frequency, amplitude and velocity over the course of a finger-tapping task [4], [5]. However, such measures sometimes show poor correlation with clinical bradykinesia rating or Parkinson's diagnosis [6], [7]. Our clinical experience is that finger-tapping amplitude in Parkinson's is sometimes observed to increase after an initial decrease in a manner that would not be captured by linear models.

We hypothesize that there may be archetypal trends in decrement that are both non-linear, and inadequately described by the current clinical scales. In this work, we use a time-series clustering approach to identify whether such archetypal trends exist and to visualize the results.

II. METHODS

A. Data Collection

We recorded videos of 137 finger-tapping examination from 69 participants at Leeds Teaching Hospitals Trust using an iPhone SE at 60 fps. 39 participants had previously diagnosed idiopathic Parkinson's disease; the remaining 30 were healthy controls. All participants had recording of both left and right hands, but one video was rejected for quality reasons. The primary unit of analysis was the number of finger-tapping examinations, as we were not concerned here with per-participant outcomes.

During the examination, participants rested their elbow on a chair arm, and the iPhone was adjusted so that hand and forearm were in frame (see Fig. 1). The lateral surface of the hand faced the camera. Participants were told to tap their thumb and index finger together as wide and as fast as possible. Each video sample was restricted to 11 seconds.

The study cohort described here is an expanded sample of those recorded previously by Wong *et al.* [8], who report additional details of the data collection method.

Videos were processed to output pixel coordinates of the participant's thumb and index fingertips at each frame using the Deeplabcut software package [9], [10]. A neurologist (SW) manually labelled 20 frames from each 660 frame video. Salient points were automatically extracted for the

[1]Zhibin Zhao is with the School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an, China zhibinzhao1993@gmail.com

[2] H. Fang is with the Department of Computer Science, Loughborough Univerity, Loughborough, UK h.fang@lboro.ac.uk

[3] S.D. Relton and S. Williams are with the Leeds Institute of Health Science, University of Leeds, UK s.d.relton@leeds.ac.uk, stefan.williams2@nhs.net

[4] J. Alty is with Leeds Teaching Hospitals NHS Trust, UK and University of Tasmania, Australia jane.alty@utas.edu.au

[6] A.J. Casson is with the Department Electrical Electronic Engineering, University of Manchester, Manchester, UK corresponding author: alex.casson@manchester.ac.uk

[6] D.C. Wong is with the Department of Computer Science and the Centre for Health Informatics, University of Manchester, Manchester, UK david.wong@manchester.ac.uk

remaining frames by the algorithm. Visual inspection of the salient points aligned to the original video were used to confirm the accuracy of the algorithm output.

22 neurologists were each assigned 30 videos at random, and asked to rate each video using the MBRS scale to provide a score between 0 and 4. Where videos were rated by multiple neurologists, the modal score was used.

B. Preprocessing

Using the pixel coordinates of the thumb and index fingertips, we generated a time series of the finger to thumb tip distance of each participant. One example is shown in Fig. 2, where outliers (large, sudden transients) in the original signal occurred due to mistakes in identifying the exact positions of fingertips by deeplabcut. To remedy this, we applied a Savitsky-Golay (SG) filter with polynomial order 3 and frame length 11 to all the time series. Fig. 2 also shows the output of the filtered time series.

C. Envelope Extraction

We then extracted the upper envelopes of the filtered signals, which represents the maximum distance between thumb and index finger tip over time. Envelope extraction consisted of finding and then interpolating the peak values.

We detected peaks using MATLAB's *findpeaks* function [11]. We restricted acceptable peak-to-peak intervals to be greater than 0.3 seconds. The threshold was determined by visual inspection, using the average interval of adjacent finger taps as a guide. To ease the undulation and maintain the wave shape, we performed shape-preserving piecewise cubic interpolation of values at neighboring peak points.

Fig. 3 shows one example of the envelope analysis in which the extracted envelope accurately reflects the variance of peak values. Finally, we normalized all the extracted envelopes through Z-score normalization using the mean and standard deviation of each envelope.

D. Clustering and Visualization of Decrement

Clustering is an unsupervised learning approach for finding intrinsic patterns in data. Typically, similarity of dataset features are compared using a distance measure, and similar features are assigned to the same cluster. Multiple approaches have been considered for time series clustering, including Gaussian process clustering [12] and Hidden Markov models [13]. Here, we implement two related methods for time series clustering based on K-means clustering.

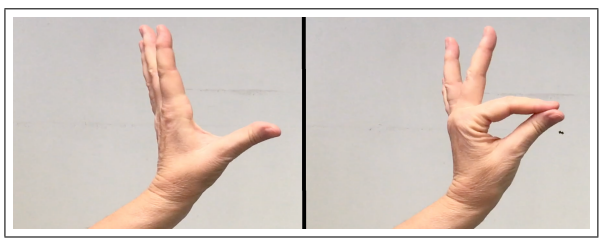

Fig. 1. Example frames from the smartphone video capture showing the hand and forearm in frame in an open (left) and closed (right) finger-tapping test position

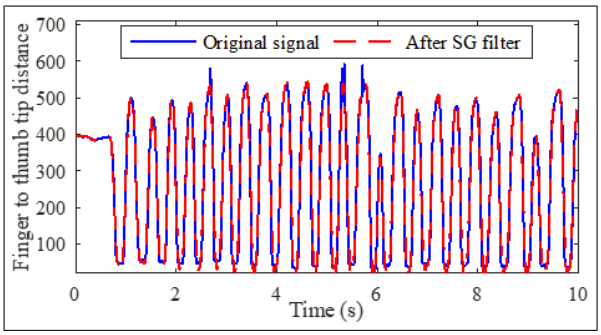

Fig. 2. One example of the original signal and the SG filtered signal

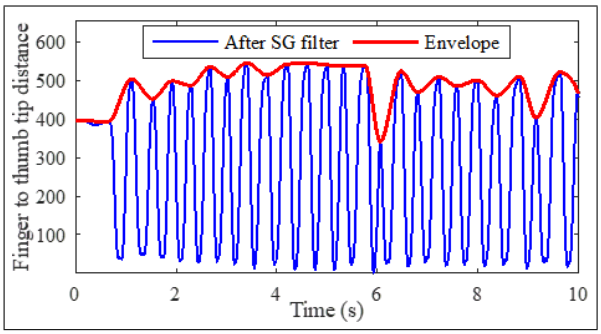

Fig. 3. One example of the envelope extraction output

The first approach was to apply K-means directly to the time series. K-means iteratively assigns a time series to the cluster centre with the lowest Euclidean distance and updates the cluster centre according to the centroid of its members. The Euclidean distance, d, between two time series, A and B, of equal length, n, is computed as:

$$d(A,B) = \sqrt{\sum_{t=1}^{n} (A_t - B_t)}$$

We selected $K = 5$ to match the number of distinct classifications within the MBRS clinical rating score. From these clusters, we sought to find the archetypal time series by extracting the arithmetic mean of the within-cluster time series. The arithmetic mean is the standard approach used to calculate an average sequence when all sequences are consistent. However, this approach risks losing shape information in instances where similar sequences are shifted in time. Fig. 4 shows one example, in which the shape of the signals in Fig. 4(a) are not preserved by the arithmetic mean in 4(b).

The second approach used an alternative distance measure, dynamic time warping (DTW). DTW is a standard measure for a large range of time series analysis applications, such as speech recognition [14], gene expression [15] and biomedical signals [16]. DTW aims to align different time sequences to find the optimal alignments. It allows for both time shifts and non-linear distortions, and does not require that the length of two sequences should be equal. The DTW measure approaches zero as two sequences become more similar.

Multiple studies have attempted to approximate an average time series [17], [18]. Among them, dynamic time warping

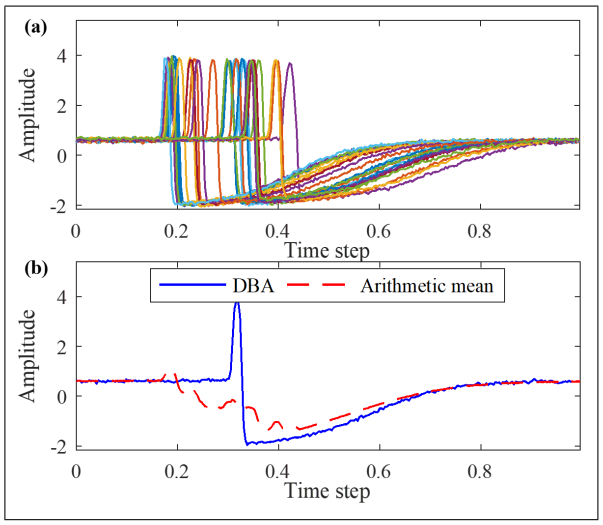

Fig. 4. One example of time series averaging. (a) the set of time series data (the first class in the Trace dataset) and (b) the average signals extracted by arithmetic mean and DBA.

barycenter averaging (DBA) proposed by Petitjean et al. is one relatively robust approach [19].

The DBA algorithm is an iterative procedure which consists of two steps: (1) Calculate the DTW measure between the temporary average sequence and each sequence in the predefined set and find the relationships (similarities) between elements of each sequence in the predefined set and elements of the temporary average sequence; (2) Update each element of the average sequence using the barycenter of elements which are related to it corresponding to the DTW measure in the above step. The convergence property, and demonstration of how DBA can be applied alongside existing clustering methods like K-means, has been proven empirically and theoretically [20].

Fig. 4(b) shows that, compared with a simple arithmetic mean, DBA preserves the local structure of a time series. We denote the clustering method used in this paper as K-DBA, signifying the use of K-means with DTW as the distance measure alongside DBA as the averaging approach.

III. RESULTS

For K-DBA, the number of centroid seeds is 10 and the number of iterations for the DBA computation is 20. The results of traditional K-means and K-DBA, in which we show the average time series (in bold red) alongside the individual time series belonging to the cluster (grey), are shown in Fig. 5 and Fig. 6, respectively. The membership of each archetypal cluster, grouped by clinically assessed MBRS decrement score, is presented in Tables I and II for K-DBA and K-means respectively.

The figures show two main time series trajectories. Fig. 5(c) and 6(a) mainly represent the participants with no bradykinesia decrement; the amplitude throughout the archetypal time series is relatively stable. As expected, the K-DBA derived signal contains greater local structure, but the two approaches provide a similar overall shape. The tables show that these clusters correspond well with those with an MBRS decrement score of 0. For 5(c) and 6(a), 37/58 and 27/38 time series were rated with an MBRS = 0, respectively.

Figs. 5(d) and 6(c) represent clusters in which decrement is clearly visible. For these clusters, the amplitude envelope decreases linearly over time. Both clusters were associated with participants assessed with more severe bradykinesia. The remaining representative clusters show large and rapid changes in amplitude at the start and end of the time series. These changes are visible in the original envelope signals but were not usually present in the raw periodic finger-tapping signal. Therefore, the pattern is most likely due to artefact introduced at the piecewise interpolation step.

TABLE I
K-MEANS CLUSTER MEMBERSHIP GROUPED BY MBRS AMPLITUDE (DECREMENT) SCORE

Cluster	MBRS Amplitude (Decrement) Score				
	0 (normal)	*1*	*2*	*3*	*4 (severe)*
(a)	8	3	3	1	1
(b)	5	2	1	4	0
(c)	37	10	8	3	0
(d)	9	9	5	7	2
(e)	4	2	4	4	1

TABLE II
K-DBA CLUSTER MEMBERSHIP GROUPED BY MBRS AMPLITUDE (DECREMENT) SCORE

Cluster	MBRS Amplitude (Decrement) Score				
	0 (normal)	*1*	*2*	*3*	*4 (severe)*
(a)	27	6	4	1	0
(b)	6	5	2	5	1
(c)	8	7	5	9	1
(d)	19	7	9	1	3
(e)	3	1	1	0	2

IV. DISCUSSION

We developed a method to visualize bradykinesia decrement directly from smartphone videos of a finger-tapping clinical assessment. We have also used a shape-preserving time series clustering approach to extract common archetypal decrement trends in a cohort of 133 videos.

Although we selected five cluster centres a priori based on the gradation of the MBRS decrement scale, we observed only two distinct time series patterns - the remaining three clusters appear to have similar characteristics. The two distinct archetypes corresponded well with participants with no bradykinesia and with high bradykinesia.

Mild grades of bradykinesia (MBRS=1-2) were not clearly associated with any cluster. It is unclear why this is the case. Mild grades of decrement might be indistinguishable from low levels (or no) decrement. However, it is more likely that the decrement patterns do not constitute an independent cluster because the data set contains relatively few examples

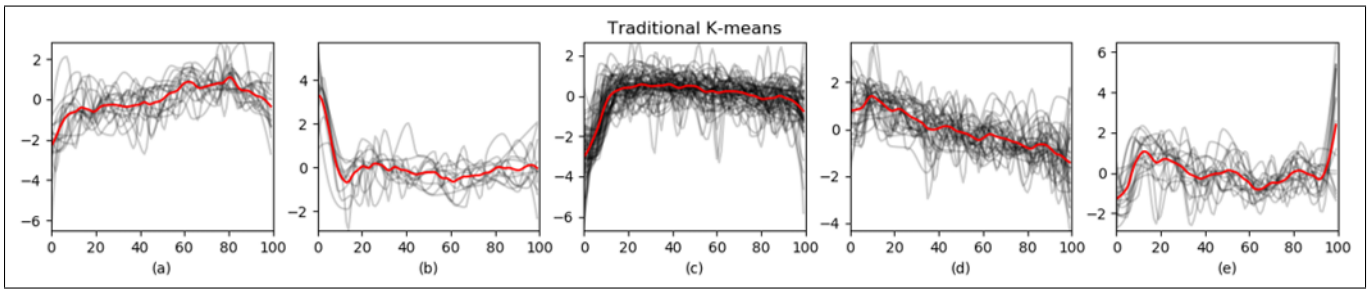

Fig. 5. The visualization results of Traditional K-means (red lines are average sequences of five class centers).

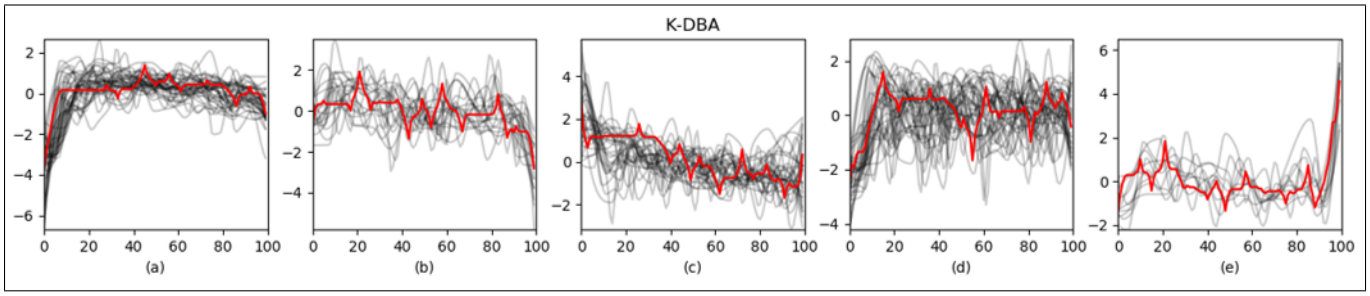

Fig. 6. The visualization results of K-DBA (red lines are average sequences of five class centers).

of mild decrement. Future work should therefore include repetition on a larger, balanced data set.

We found no evidence to lend support to the initial hypothesis, that there exists non-linear *decrement* trajectories. This means that current quantitative metrics using linear regression are likely to be sufficient, and that previous studies can be interpreted as evidence that some bradykinesia does not involve decrement [7], rather than reflecting measurement limitations. The method shown here is flexible and extensible to other clinical time series data. It allows grouping of similar time series trajectories to identify common trend patterns.

REFERENCES

[1] R.B. Postuma, D. Berg, M. Stern, W. Poewe, C.W. Olanow, W. Oertel *et al.* "MDS clinical diagnostic criteria for Parkinson's disease. Movement Disorders. Vol. 30, no. 12, pp. 1591-1601, 2015.

[2] P. Martinez-Martin, A. Gil-Nagel, L. Morlan Gracia, J. Balserio Gomez, J. Martines-Sarries, F. Bermejo. "Unified Parkinson's disease rating scale characteristics and structure," Movement Disorders. Vol. 9, no. 1, pp. 76–83, 1994.

[3] D. Heldman, J.P. Giuffrida, R. Chen, M. Payne, F. Mazzella, A.P. Duker *et al.* "The modified bradykinesia rating scale for Parkinson's disease: Reliability and comparison with kinematic measures," Movement Disorders, vol. 26, no. 10, pp 1859–1863, 2011.

[4] P.J. Bank, J. Marinus, C.G. Meskers, J.H. de Groot, J.J. van Hilten. "Optical Hand Tracking: A Novel Technique for the Assessment of Bradykinesia in Parkinson's Disease," Movement disorders clinical practice, vol. 4, no. 6, pp. 875-83, 2017

[5] O. Martinez-Manzanera, E. Roosma, M. Beudel, R.W. Borgemeester, T. van Laar, N.M. Maurits. "A method for automatic and objective scoring of bradykinesia using orientation sensors and classification algorithms," IEEE Trans. on Biomed. Eng. Vol. 63, No. 5, pp.1016-24, 2015.

[6] E. Ruzicka, R. Krupicka, K. Zarubova, J. Rusz, R. Jech, Z. Szabo. "Tests of manual dexterity and speed in Parkinson's disease: Not all measure the same," Parkinsonism Related Disorders, vol. 28, pp. 118–123, 2016

[7] M. Bologna, G. Leodori, P. Stirpe, G. Paparella, D. Colella, D. Belvisi, *et al.* "Bradykinesia in early and advanced Parkinson's disease", J. Neurol. Sci., vol. 369, pp. 286–291, 2016.

[8] D.C Wong, S.D. Relton, H. Fang, R. Qhawaji, C.D. Graham, J. Alty, S. Williams. "Supervised classification of bradykinesia for Parkinson's disease diagnosis from smartphone videos," Proc. IEEE 32nd Int Symp on Computer-Based Medical Systems (CBMS) 2019 Mar 29.

[9] A. Mathis *et al.* "DeepLabCut: markerless pose estimation of user-defined body parts with deep learning," Nature Neuroscience, vol 21, pp. 1281-1289, 2018.

[10] S. Williams, Z. Zhao, A. Hafeez, D.C. Wong, S.D. Relton, H. Fang, J. Alty. "The discerning eye of computer vision: can it measure Parkinson's finger tap bradykinesia," J. Neurol. Sci., submitted for publication.

[11] MATLAB and Statistics Toolbox Release 2018a, The Mathworks, Inc., Natick, Massachusetts, United States.

[12] M.A. Pimentel, D.A. Clifton, L. Tarassenko. "Gaussian process clustering for the functional characterisation of vital-sign trajectories," Proc. - 2013 IEEE Int. Workshop on Machine Learning for Signal Process. (MLSP) pp. 1-6, 2013.

[13] E.B. Fox, E.B. Sudderth, M.I. Jordan, A.S. Willsky, "Sharing Features among Dynamical Systems with Beta Processes," Proc. - Neural Information Processing Systems, Vancouver, Canada December 2009.

[14] H. Sakoe "A dynamic programming approach to continuous speech recognition," Proc. – Int. Congress of Acoustics, Budapest, Hungary, 1971.

[15] J. Aach, G.M. Church. "Aligning gene expression time series with time warping algorithms," Bioinformatics, vol. 17, no. 6, pp. 495-508, 2001.

[16] D. Wong, I Strachan, L Tarassenko. "Visualisation of high-dimensional data for very large data sets," Proc. - Workshop Mach. Learn. Healthcare Appl., 25th Int. Conf. Mach. Learn., Helsinki, Finland, 2008.

[17] L. Gupta, D.L. Molfese, R. Tammana, P.G. Simos. "Nonlinear alignment and averaging for estimating the evoked potential," IEEE Trans. Biomed. Eng, vol. 43, no.4, pp.348-56, 1996.

[18] V. Niennattrakul, C.A. Ratanamahatana. "Inaccuracies of shape averaging method using dynamic time warping for time series data," Proc. – Int. Conf. Computational Science, Berlin, May 2007.

[19] F. Petitjean, A. Ketterlin, P. Gançarski. "A global averaging method for dynamic time warping, with applications to clustering," Pattern Recognition, vol. 44, no.3 pp. 678-693, 2011.

[20] F. Petitjean, G. Forestier, G.I. Webb, A.E. Nicholson, Y. Chen, E. Keogh. "Dynamic time warping averaging of time series allows faster and more accurate classification," Proc. – Int. Conf. Data Mining, pp. 470-479, 2014.